#4

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/827,937A

DATE: 08/15/2001
TIME: 14:22:49

Input Set : A:\seqlist_1488 1220003.txt
Output Set: N:\CRF3\08152001\I827937A.raw

ENTERED

```
  5 <110> APPLICANT: Li, Yi
  8        Ruben, Steven M.
 13 <120> TITLE OF INVENTION: Human G-Protein Coupled Receptors
 19 <130> FILE REFERENCE: 1488.1220003
 25 <140> CURRENT APPLICATION NUMBER: 09/827,937A
 28 <141> CURRENT FILING DATE: 2001-04-09
 33 <150> PRIOR APPLICATION NUMBER: 08/852,824
 36 <151> PRIOR FILING DATE: 1997-05-07
 41 <160> NUMBER OF SEQ ID NOS: 18
 47 <170> SOFTWARE: PatentIn Ver. 2.0
 53 <210> SEQ ID NO: 1
 56 <211> LENGTH: 2247
 59 <212> TYPE: DNA
 62 <213> ORGANISM: Homo sapiens
 66 <220> FEATURE:
 69 <221> NAME/KEY: CDS
 72 <222> LOCATION: (226)..(1251)
 78 <400> SEQUENCE: 1
 81 gcacgaggaa cagaacactt tctcatgtcc agggtcagat tacaagagca ctcaagactt 60
 87 tactgacgaa aactcaggaa atcctctatc acaaagaggt ttggcaacta aactaagaca 120
 93 ttaaaaggaa aataccagat gccactctgc aggctgcaat aactactact tactggatac 180
 99 attcaaaccc tccagaatca acagttatca ggtaaccaac aagaa atg caa gcc gtc 237
                                                   Met Gln Ala Val
102                                                  1
105
111 gac aat ctc acc tct gcg cct ggg aac acc agt ctg tgc acc aga gac 285
114 Asp Asn Leu Thr Ser Ala Pro Gly Asn Thr Ser Leu Cys Thr Arg Asp
117  5                    10                   15                   20
123 tac aaa atc acc cag gtc ctc ttc cca ctg ctc tac act gtc ctg ttt 333
126 Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Leu Tyr Thr Val Leu Phe
129                   25                   30                   35
135 ttt gtt gga ctt atc aca aat ggc ctg gcg atg agg att ttc ttt caa 381
138 Phe Val Gly Leu Ile Thr Asn Gly Leu Ala Met Arg Ile Phe Phe Gln
141                   40                   45                   50
147 atc cgg agt aaa tca aac ttt att att ttt ctt aag aac aca gtc att 429
150 Ile Arg Ser Lys Ser Asn Phe Ile Ile Phe Leu Lys Asn Thr Val Ile
153              55                   60                   65
159 tct gat ctt ctc atg att ctg act ttt cca ttc aaa att ctt agt gat 477
162 Ser Asp Leu Leu Met Ile Leu Thr Phe Pro Phe Lys Ile Leu Ser Asp
165              70                   75                   80
171 gcc aaa ctg gga aca gga cca ctg aga act ttt gtg tgt caa gtt acc 525
174 Ala Lys Leu Gly Thr Gly Pro Leu Arg Thr Phe Val Cys Gln Val Thr
177 85                   90                   95                   100
183 tcc gtc ata ttt tat ttc aca atg tat atc agt att tca ttc ctg gga 573
186 Ser Val Ile Phe Tyr Phe Thr Met Tyr Ile Ser Ile Ser Phe Leu Gly
189                   105                  110                  115
195 ctg ata act atc gat cgc tac cag aag acc acc agg cca ttt aaa aca 621
198 Leu Ile Thr Ile Asp Arg Tyr Gln Lys Thr Thr Arg Pro Phe Lys Thr
```

```
                            120                  125                  130
201
207 tcc aac ccc aaa aat ctc ttg ggg gct aag att ctc tct gtt gtc atc   669
210 Ser Asn Pro Lys Asn Leu Leu Gly Ala Lys Ile Leu Ser Val Val Ile
213         135                  140                  145
219 tgg gca ttc atg ttc tta ctc tct ttg cct aac atg att ctg acc aac   717
222 Trp Ala Phe Met Phe Leu Leu Ser Leu Pro Asn Met Ile Leu Thr Asn
225         150                  155                  160
231 agg cag ccg aga gac aag aat gtg aag aaa tgc tct ttc ctt aaa tca   765
234 Arg Gln Pro Arg Asp Lys Asn Val Lys Lys Cys Ser Phe Leu Lys Ser
237 165                  170                  175                  180
243 gag ttc ggt cta gtc tgg cat gaa ata gta aat tac atc tgt caa gtc   813
246 Glu Phe Gly Leu Val Trp His Glu Ile Val Asn Tyr Ile Cys Gln Val
249                  185                  190                  195
255 att ttc tgg att aat ttc tta att gtt att gta tgt tat aca ctc att   861
258 Ile Phe Trp Ile Asn Phe Leu Ile Val Ile Val Cys Tyr Thr Leu Ile
261         200                  205                  210
267 aca aaa gaa ctg tac cgg tca tac gta aga acg agg ggt gta ggt aaa   909
270 Thr Lys Glu Leu Tyr Arg Ser Tyr Val Arg Thr Arg Gly Val Gly Lys
273         215                  220                  225
279 gtc ccc agg aaa aag gtg aac gtc aaa gtt ttc att atc att gct gta   957
282 Val Pro Arg Lys Lys Val Asn Val Lys Val Phe Ile Ile Ile Ala Val
285         230                  235                  240
291 ttc ttt att tgt ttt gtt cct ttc cat ttt gcc cga att cct tac acc  1005
294 Phe Phe Ile Cys Phe Val Pro Phe His Phe Ala Arg Ile Pro Tyr Thr
297 245                  250                  255                  260
303 ctg agc caa acc cgg gat gtc ttt gac tgc act gct gaa aat act ctg  1053
306 Leu Ser Gln Thr Arg Asp Val Phe Asp Cys Thr Ala Glu Asn Thr Leu
309                  265                  270                  275
315 ttc tat gtg aaa gag agc act ctg tgg tta act tcc tta aat gca tgc  1101
318 Phe Tyr Val Lys Glu Ser Thr Leu Trp Leu Thr Ser Leu Asn Ala Cys
321         280                  285                  290
327 ctg gat ccg ttc atc tat ttt ttc ctt tgc aag tcc ttc aga aat tcc  1149
330 Leu Asp Pro Phe Ile Tyr Phe Phe Leu Cys Lys Ser Phe Arg Asn Ser
333         295                  300                  305
339 ttg ata agt atg ctg aag tgc ccc aat tct gca aca tct ctg tcc cag  1197
342 Leu Ile Ser Met Leu Lys Cys Pro Asn Ser Ala Thr Ser Leu Ser Gln
345         310                  315                  320
351 gac aat agg aaa aaa gaa cag gat ggt ggt gac cca aat gaa gag act  1245
354 Asp Asn Arg Lys Lys Glu Gln Asp Gly Gly Asp Pro Asn Glu Glu Thr
357 325                  330                  335                  340
363 cca atg taaacaaatt aactaaggaa atatttcaat ctctttgtgt tcagaactcg  1301
366 Pro Met
372 ttaaagcaaa gcgctaagta aaaatattaa ctgacgaaga agcaactaag ttaataataa 1361
378 tgactctaaa gaaacagaag attacaaaag caattttcat ttcctttcc agtatgaaaa 1421
384 gctatcttaa aatatagaaa actaatctaa actgtagctg tattagcagc aaaacaaacg 1481
390 acatccaatt gtcatgctgc atgcaaaact acacagaatt catgttttgg cagagttttg 1541
396 gcaaaatgag taatcatata atatttactg taatttttaa aatacattat cgttcacaat 1601
402 tttatttttt cataatcaac taaggaagaa cgatcaattg gatataatct tcttaccaaa 1661
408 aatgatagtt aaaatgtata tatatcctag tcccctaacc aaatcctgac ctattgggat 1721
```

```
414 acttataaaa atttaagtaa gtgggataca caaagaataa taactattaa cttttcatta 1781
420 ttagccaaaa acctaaggga tttaaactaa ttgaaactgt atttgattgg acttaatttt 1841
426 ttatgtttat ttagaagata aagatttaag aagaccttta caataaagag aagaaatatc 1901
432 gaagtcatta aaataaggag acttactttt atgacattct aatactaaaa aatatagaaa 1961
438 tatttcctta attctagaga aactagtttt actaattttt tacaacttca ataataccat 2021
444 cactgacact tacctttatt aattagcttc tagaaaatag ctgctaatta ggttaatgaa 2081
450 cattttacct tagtgaaaaa aaattaatta aatatgatta caaagttgca cagcataact 2141
456 actgagagga aagtgattga tctgtttgta attacttgtt tgtattggtg tgtataaaat 2201
462 acaaatttac attaaactct aaatcattaa aaaaaaaaaa aaaaaa 2247

471 <210> SEQ ID NO: 2
474 <211> LENGTH: 342
477 <212> TYPE: PRT
480 <213> ORGANISM: Homo sapiens
486 <400> SEQUENCE: 2
489 Met Gln Ala Val Asp Asn Leu Thr Ser Ala Pro Gly Asn Thr Ser Leu
492 1               5                   10                  15
498 Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Leu Tyr
501             20                  25                  30
507 Thr Val Leu Phe Phe Val Gly Leu Ile Thr Asn Gly Leu Ala Met Arg
510         35                  40                  45
516 Ile Phe Phe Gln Ile Arg Ser Lys Ser Asn Phe Ile Ile Phe Leu Lys
519     50                  55                  60
525 Asn Thr Val Ile Ser Asp Leu Leu Met Ile Leu Thr Phe Pro Phe Lys
528 65                  70                  75                  80
534 Ile Leu Ser Asp Ala Lys Leu Gly Thr Gly Pro Leu Arg Thr Phe Val
537             85                  90                  95
543 Cys Gln Val Thr Ser Val Ile Phe Tyr Phe Thr Met Tyr Ile Ser Ile
546             100                 105                 110
552 Ser Phe Leu Gly Leu Ile Thr Ile Asp Arg Tyr Gln Lys Thr Thr Arg
555         115                 120                 125
561 Pro Phe Lys Thr Ser Asn Pro Lys Asn Leu Leu Gly Ala Lys Ile Leu
564     130                 135                 140
570 Ser Val Val Ile Trp Ala Phe Met Phe Leu Leu Ser Leu Pro Asn Met
573 145                 150                 155                 160
579 Ile Leu Thr Asn Arg Gln Pro Arg Asp Lys Asn Val Lys Lys Cys Ser
582             165                 170                 175
588 Phe Leu Lys Ser Glu Phe Gly Leu Val Trp His Glu Ile Val Asn Tyr
591             180                 185                 190
597 Ile Cys Gln Val Ile Phe Trp Ile Asn Phe Leu Ile Val Ile Val Cys
600             195                 200                 205
606 Tyr Thr Leu Ile Thr Lys Glu Leu Tyr Arg Ser Tyr Val Arg Thr Arg
609     210                 215                 220
615 Gly Val Gly Lys Val Pro Arg Lys Lys Val Asn Val Lys Val Phe Ile
618 225                 230                 235                 240
624 Ile Ile Ala Val Phe Phe Ile Cys Phe Val Pro Phe His Phe Ala Arg
627             245                 250                 255
633 Ile Pro Tyr Thr Leu Ser Gln Thr Arg Asp Val Phe Asp Cys Thr Ala
636             260                 265                 270
642 Glu Asn Thr Leu Phe Tyr Val Lys Glu Ser Thr Leu Trp Leu Thr Ser
```

```
                             280                   285
645          275        Phe Ile Tyr Phe Phe Leu Cys Lys Ser
651 Leu Asn Ala Cys Leu Asp Pro            300
654     290                    295
660 Phe Arg Asn Ser Leu Ile Ser Met Leu Lys Cys Pro Asn Ser Ala Thr
663 305              310                 315                   320
669 Ser Leu Ser Gln Asp Asn Arg Lys Lys Glu Gln Asp Gly Gly Asp Pro
672              325               330                   335
678 Asn Glu Glu Thr Pro Met
681              340
690 <210> SEQ ID NO: 3
693 <211> LENGTH: 1637
696 <212> TYPE: DNA
699 <213> ORGANISM: Homo sapiens
705 <220> FEATURE:
708 <221> NAME/KEY: CDS
711 <222> LOCATION: (50)..(1201)
717 <400> SEQUENCE: 3
720 ggcacgagcc caccctgcgt cgggcctcag tcagcccccg ggggaggcc atg aac gcc  58
723                                                       Met Asn Ala
726                                                         1
732 acg ggg acc ccg gtg gcc ccc gag tcc tgc caa cag ctg gcg gcc ggc 106
735 Thr Gly Thr Pro Val Ala Pro Glu Ser Cys Gln Gln Leu Ala Ala Gly
738     5                   10                  15
744 ggg cac agc cgg ctc att gtt ctg cac tac aac cac tcg ggc cgg ctg 154
747 Gly His Ser Arg Leu Ile Val Leu His Tyr Asn His Ser Gly Arg Leu
750 20                  25                  30                  35
756 gcc ggg cgc ggg ggg ccg gag gat ggc ggc ctg ggg gcc ctg cgg ggg 202
759 Ala Gly Arg Gly Gly Pro Glu Asp Gly Gly Leu Gly Ala Leu Arg Gly
762              40                  45                  50
768 ctg tcg gtg gcc gcc agc tgc ctg gtg gtg ctg gag aac ttg ctg gtg 250
771 Leu Ser Val Ala Ala Ser Cys Leu Val Val Leu Glu Asn Leu Leu Val
774              55                  60                  65
780 ctg gcg gcc atc acc agc cac atg cgg tcg caa cgc tgg gtc tac tat 298
783 Leu Ala Ala Ile Thr Ser His Met Arg Ser Gln Arg Trp Val Tyr Tyr
786          70                  75                  80
792 tgc ctg gtg aac att acg atg agt gac ctg ctc acg ggc gcg gcc tac 346
795 Cys Leu Val Asn Ile Thr Met Ser Asp Leu Leu Thr Gly Ala Ala Tyr
798          85                  90                  95
804 ctg gcc aac gtg ctg ctg tcg ggg gcc cgc acc ttc cgt ctg gcg ccc 394
807 Leu Ala Asn Val Leu Leu Ser Gly Ala Arg Thr Phe Arg Leu Ala Pro
810 100                 105                 110                 115
816 gcc cag tgg ttc cta cgg aag ggc ctg ctc ttc acc gcc ctg gcc gcc 442
819 Ala Gln Trp Phe Leu Arg Lys Gly Leu Leu Phe Thr Ala Leu Ala Ala
822              120                 125                 130
828 tcc acc ttc agc ctg ctc ttc act gca ggg ttg cgc ttt gcc acc atg 490
831 Ser Thr Phe Ser Leu Leu Phe Thr Ala Gly Leu Arg Phe Ala Thr Met
834              135                 140                 145
840 gtg cgg ccg gtg gcc gag agc ggg gcc acc aag acc agc cgc gtc tac 538
843 Val Arg Pro Val Ala Glu Ser Gly Ala Thr Lys Thr Ser Arg Val Tyr
```

```
846          150                        155                    160
852 ggc ttc atc ggc ctc tgc tgg ctg ctg gcc gcg ctg ctg ggg atg ctg    586
855 Gly Phe Ile Gly Leu Cys Trp Leu Leu Ala Ala Leu Leu Gly Met Leu
858     165                        170                    175
864 cct ttg ctg ggc tgg aac tgc ctg tgc gcc ttt gac cgc tgc tcc agc    634
867 Pro Leu Leu Gly Trp Asn Cys Leu Cys Ala Phe Asp Arg Cys Ser Ser
870 180                    185                    190              195
876 ctt ctg ccc ctc tac tcc aag cgc tac atc ctc ttc tgc ctg gtg atc    682
879 Leu Leu Pro Leu Tyr Ser Lys Arg Tyr Ile Leu Phe Cys Leu Val Ile
882                    200                    205              210
888 ttc gcc ggc gtc ctg gcc acc atc atg ggc ctc tat ggg gcc atc ttc    730
891 Phe Ala Gly Val Leu Ala Thr Ile Met Gly Leu Tyr Gly Ala Ile Phe
894             215                    220                    225
900 cgc ctg gtg cag gcc agc ggg cag aag gcc cca cgc cca gcg gcc cgc    778
903 Arg Leu Val Gln Ala Ser Gly Gln Lys Ala Pro Arg Pro Ala Ala Arg
906         230                    235                    240
912 cgc aag gcc cgc cgc ctg ctg aag acg gtg ctg atg atc ctg ctg gcc    826
915 Arg Lys Ala Arg Arg Leu Leu Lys Thr Val Leu Met Ile Leu Leu Ala
918     245                    250                    255
924 ttc ttg gtg tgc tgg gga cca ctc ttc ggg ctg ctg ctg gcc gac gtc    874
927 Phe Leu Val Cys Trp Gly Pro Leu Phe Gly Leu Leu Leu Ala Asp Val
930 260                    265                    270              275
936 ttt ggc tcc aac ctc tgg gcc cag gag tac ctg cgg ggc atg gac tgg    922
939 Phe Gly Ser Asn Leu Trp Ala Gln Glu Tyr Leu Arg Gly Met Asp Trp
942                    280                    285              290
948 atc ctg gcc ctg gcc gtc ctc aac tcg gcg gtc aac ccc atc atc tac    970
951 Ile Leu Ala Leu Ala Val Leu Asn Ser Ala Val Asn Pro Ile Ile Tyr
954             295                    300                    305
960 tcc ttc cgc agc agg gag gtg tgc aga gcc gtg ctc agc ttc ctc tgc   1018
963 Ser Phe Arg Ser Arg Glu Val Cys Arg Ala Val Leu Ser Phe Leu Cys
966         310                    315                    320
972 tgc ggg tgt ctc cgg ctg ggc atg cga ggg ccc ggg gac tgc ctg gcc   1066
975 Cys Gly Cys Leu Arg Leu Gly Met Arg Gly Pro Gly Asp Cys Leu Ala
978     325                    330                    335
984 cgg gcc gtc gag gct cac tcc gga gct tcc acc acc gac agc tct ctg   1114
987 Arg Ala Val Glu Ala His Ser Gly Ala Ser Thr Thr Asp Ser Ser Leu
990 340                    345                    350              355
996 agg cca agg gac agc ttt cgc ggc tcc cgc tcg ctc agc ttt cgg atg   1162
999 Arg Pro Arg Asp Ser Phe Arg Gly Ser Arg Ser Leu Ser Phe Arg Met
1002             360                    365                    370
1008 cgg gag ccc ctg tcc agc atc tcc agc gtg cgg agc atc tgaagttgca  1211
1011 Arg Glu Pro Leu Ser Ser Ile Ser Ser Val Arg Ser Ile
1014                 375                    380
1020 gtcttgcgtg tggatggtgc aaccaccggg tgcgtgccag gcaggccctc ctggggtaca 1271
1026 ggaagctgtg tgcacgcaac ctcgccctgt atggggagca gggaacggga caggcccca 1331
1032 tggacttgcc cggtggcctc tcggggcttc tgacgccata tggacttgcc cattgcctat 1391
1038 ggctcaccct ggacaaggag gcaaccaccc cacctccccg taggagcaga gagcaccctg 1451
1044 gtgtgggggc gagtgggttc cccacaaccc cgcttctgtg tgattctggg gaagtcccgg 1511
1050 cccctctctg ggcctcagta gggctcccag gctgcaaggg gtggactgtg ggatgcatgc 1571
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/827,937A

DATE: 08/15/2001
TIME: 14:22:50

Input Set : **A:\seqlist_1488 1220003.txt**
Output Set: **N:\CRF3\08152001\I827937A.raw**

```
L:1317 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:1353 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:1390 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:1426 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:1462 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:1498 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:1534 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:1570 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:1606 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:1642 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:1678 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:1714 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
```